# R. Clifton Bailey Statistics Seminar Series

# Density estimation in infinite dimensional exponential families


## Bharath Sriperumbudur

### Department of Statistics
### Penn State University


Johnson Center G19 – Gold Room
4000 University Drive, Fairfax, VA 22030
March 23, 2018
11:00 A.M. - 12:00 Noon


**Abstract: We consider an infinite dimensional generalization of natural exponential family of probability densities, which are parametrized by functions in a reproducing kernel Hilbert space (RKHS), and show it to be quite rich in the sense that a broad class of densities on R^d can be approximated arbitrarily well in Kullback-Leibler (KL) divergence by elements in the infinite dimensional family, P. Motivated by this approximation property, we consider the problem of estimating an unknown density p_0, through an element in P. Standard techniques like maximum likelihood estimation (MLE) or pseudo MLE (based on the method of sieves), which are based on minimizing the KL divergence between p_0 and P, do not yield practically useful estimators because of their inability to efficiently handle the log-partition function. We propose an estimator based on minimizing the Fisher divergence between p_0 and P, which involves solving a simple finite-dimensional linear system. We show the proposed estimator to be consistent, and**

**provide convergence rates under a smoothness assumption that $\log(p_0)$ belongs to the image of the fractional power of a Hilbert-Schmidt operator defined on RKHS. Through numerical simulations we demonstrate that the proposed estimator outperforms the non-parametric kernel density estimator, and that the advantage of the proposed estimator grows with increasing dimension.**